

(TM)

Release 3.1A John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

MPsrch_pp protein - protein database search, using Smith-Waterman algorithm

Run on: Mon Oct 5 10:51:34 1998; MasPar time 16.33 Seconds
587.549 Million cell updates/sec

Tabular output not generated.

Title: >US-08-908-884-3
Description: (1-593) from US08908884.pep
Perfect Score: 4122
Sequence: 1 MDTTIDGFADSYEISSTSFSV..... STSKSTGGKRSNRKLSHRRR 593

Scoring table: PAM 150
Gap 11

Searched: 131922 seqs, 16180660 residues

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: a-geneseq32
1:part1 2:part2 3:part3 4:part4 5:part5 6:part6 7:part7
8:part8 9:part9 10:part10 11:part11 12:part12 13:part13
14:part14 15:part15 16:part16 17:part17 18:part18
19:part19 20:part20 21:part21 22:part22 23:part23
24:part24 25:part25 26:part26 27:part27 28:part28
29:part29

Statistics: Mean 36.233; Variance 181.818; scale 0.199

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	%	Description	Pred. No.
						Query		

1	4122	100.0	593	29	W23963	Arabidopsis thaliana	0.00e+00
2	180	4.4	991	16	R80096	Black widow spider de	1.80e-04
3	180	4.4	1214	16	R80097	Black widow spider de	1.80e-04
4	125	3.0	679	14	R82661	Partial murine 2-5A-d	1.17e+00
5	125	3.0	679	21	W12703	Mouse 2-5A-dependent	1.17e+00
6	125	3.0	679	11	R59077	2-5A-dependent RNA-as	1.17e+00
7	123	3.0	741	14	R82660	Human 2-5A-dependent	1.58e+00
8	123	3.0	741	11	R59076	2-5A-dependent RNA-as	1.58e+00
9	123	3.0	741	14	R82659	Human 2-5A-dependent	1.58e+00
10	123	3.0	741	21	W12702	Human 2-5A-dependent	1.58e+00
11	115	2.8	339	17	R89748	AFT-1 interacting pro	5.24e+00
12	113	2.7	226	21	W15483	Human P28.	7.04e+00
13	110	2.7	718	22	W01537	Bovine ribonuclease L	1.09e+01
14	106	2.6	281	19	W02622	Yeast GCN4 transcript	1.95e+01
15	108	2.6	381	21	W15462	Haemangioma factor-1.	1.46e+01
16	107	2.6	587	8	R41889	Thermostable pyruvic	1.69e+01
17	108	2.6	1078	6	R28963	Notch hN3k full lengt	1.46e+01
18	104	2.5	281	21	W13955	LZ variant of GCN4.	2.60e+01
19	103	2.5	403	22	W11218	Leishmania braziliens	3.00e+01
20	102	2.5	403	23	W24290	LbeIF4A.	3.45e+01
21	102	2.5	403	15	R77503	Leishmania sp. antige	3.45e+01
22	102	2.5	403	23	W24291	LmeIF4A.	3.45e+01
23	103	2.5	422	11	R58599	Fowlpox virus protein	3.00e+01
24	103	2.5	422	11	R58855	Fowlpox virus (FPV) g	3.00e+01
25	103	2.5	422	21	W10693	Homology vector 443-8	3.00e+01
26	102	2.5	440	17	R94600	S4 protein.	3.45e+01
27	101	2.5	752	17	R83018	Calcium-independent c	3.98e+01
28	101	2.5	752	22	W17849	Hamster cytosolic pho	3.98e+01
29	101	2.5	752	20	W01479	Calcium-independent c	3.98e+01
30	101	2.5	752	21	W13163	Ca-independent phosph	3.98e+01
31	99	2.4	260	15	R85485	Murine CD40 ligand.	5.28e+01
32	98	2.4	260	7	R36702	Murine CD40-L.	6.07e+01
33	97	2.4	260	28	W41179	CD40 ligand.	6.98e+01
34	97	2.4	260	10	R53970	Mouse CD40-L type II	6.98e+01
35	97	2.4	281	19	W02621	Yeast GCN4 transcript	6.98e+01
36	98	2.4	293	13	R74644	Smallpox virus strain	6.07e+01
37	97	2.4	317	6	R29577	IkB NF-kappa-B-bindin	6.98e+01
38	100	2.4	365	1	P91461	Rpt-1 clone cDNA inse	4.58e+01
39	97	2.4	560	28	W26763	Human DNA replication	6.98e+01
40	97	2.4	732	23	W14068	Helicobacter pylori f	6.98e+01
41	98	2.4	976	29	W41377	Rat protein p138.	6.07e+01
42	98	2.4	1030	29	W41378	Human protein p164.	6.07e+01
43	99	2.4	1052	15	R88577	Mouse focal adhesion	5.28e+01
44	98	2.4	1052	15	R88576	Human focal adhesion	6.07e+01
45	99	2.4	2185	2	R12141	Enteroviral polypepti	5.28e+01

ALIGNMENTS

RESULT 1
ID W23963 standard; Protein; 593 AA.
AC W23963;
DT 17-JUN-1998 (first entry)

DE Arabidopsis thaliana NIM1 protein.
KW NIM1; noninducible immunity; disease resistance; plants;
KW SAR gene expression.
OS Arabidopsis thaliana.
PN WO9749822-A1.
PD 31-DEC-1997.
PF 10-MAR-1997; E01218.
PR 10-JAN-1997; US-035022.
PR 21-JUN-1996; US-020272.
PR 30-AUG-1996; US-024883.
PR 13-DEC-1996; US-033177.
PR 27-DEC-1996; US-773559.
PA (NOVS) NOVARTIS AG.
PI Delaney TP, Ellis DM, Friedrich LB, Johnson JE,
PI Lawton KA, Ryals JA, Weymann K;
DR WPI; 98-077185/07.
DR N-PSDB; V04632.
PT NIM1 gene which allows activation in plant of systemic acquired
PT resistance - useful to confer broad spectrum disease resistance in
PT plants, specifically crop plants, e.g. rice, wheat, barley, rye and
PT corn
PS Disclosure; Fig 15; 153pp; English.
CC The sequence is that of the NIM1 (noninducible immunity) protein. It
CC may be used to confer a broad spectrum disease resistance in plants,
CC specifically crop plants, e.g. rice, wheat, barley, rye and corn.
CC The NIM1 gene can be used to confer universal disease susceptibility
CC to plant cells, and their progeny. It can also be used in a screening
CC method for identifying compounds capable of inducing broad spectrum
CC disease resistance in plants, while the plant cells, and their
CC progeny, can be used to isolate a gene fragment which allows
CC expression of broad spectrum disease resistance in plants, or to
CC incorporate the resistant trait into plant lines through breeding.
SQ Sequence 593 AA;

Query Match 100.0%; Score 4122; DB 29; Length 593;
Best Local Similarity 100.0%; Pred. No. 0.00e+00;
Matches 593; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

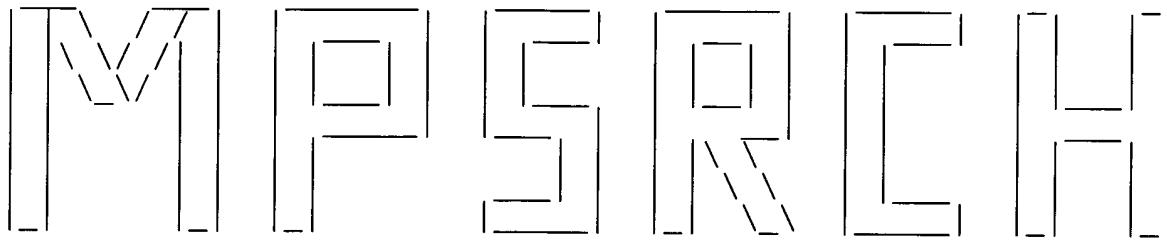
Db 1 mdttidgfadsyeisstsfvatdntdssivylaeeqvlgtgpdsalqlsnsfesvfdsp 60
Qy 1 MDTTIDGFADSYEISSTSFVATDNTDSSIVYLAEEQVLGTGPDSALQLLSNSFESVFDSP 60

Db 61 ddfysdaklvlsdgrevsfhrcvlsarssfksalaaakkedsnnntaavklelkeakd 120
Qy 61 DDFYSDAKLVLSDGREVSFHRCVLSARSSFFKSALAAAKKEKDSNNNTAAVKLELKEIAD 120

Db 121 yevgfdsvvtvlavyssrvrpppkvgsecadencchvacrpavdfmlevlylafifkip 180
Qy 121 YEVGFDSVVTVLAYVYSSRVRPPPKGVSECADENCCHVACRPAVDFMLEVLYLAFIFKIP 180

Db 181 elitlyqrh1ldvvdkvviedtlvilklanicgkacmklldrckeiiivksnvdmsleks 240
Qy 181 ELITLYQRHLLDVVDKVVIEDTLVILKLANICGKACMKLLDRCKEIIIVKSNVDMSLEKS 240

Db	241	lppeelvkeiidrrkelglevpkvvkhvsvnhkalddielvklllkedhtnlldacalh	300
Qy	241	LPEELVKEIIDRRKELGLEVPKVKHVSNVHKALDSDDIELVKLLLKEDHTNLDDACALH	300
Db	301	favaycnvktatdllkldladvnhrnprgytvlhvaamrkepqlilsllekgasaseatl	360
Qy	301	FAVAYCNVKTATDLLKLDIADVNHRNPRGYTVLHVAAMRKEPQLILSLEKGASASEATL	360
Db	361	egrtalmiakqatmavecnnipeqckhs1kgrlcveileqedkreqiprdvppsfavaad	420
Qy	361	EGRTALMIAKQATMAVECNNIPEQCKHSLKGRLCVEILEQEDKREQIPRDVPPSFAVAAD	420
Db	421	elkmtdlenrvvalaqrlfpteaqaaameiaemkgtcefivtslepdrltgtkrtspgvk	480
Qy	421	ELKMTLLDLENRVALAQRLFPTEAQAAAMEIAEMKGTCEFIVTSLEPDRLTGTKRTSPGVK	480
Db	481	iapfrileehqsrlkalsktvelgkrffprcsavldqimncedltqlacgeddtækrlq	540
Qy	481	IAPFRILEEHQSRLKALKSFTVELGKRFFPRCSAVLDQIMNCEDLTQLACGEDDTAEKRLQ	540
Db	541	kkqrymeiqetlkkafsednlelgnsstddstsstsksstggkrsnrklshrrr	593
Qy	541	KKQRYMEIQETLKAFSEDNLELGNSSTDDSTSSTSSTGGKRSNRKLSHRRR	593



(TM)

Release 3.1A John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

MPsrch_nn n.a. - n.a. database search, using Smith-Waterman algorithm

Run on: Mon Oct 5 20:40:04 1998; MasPar time 2645.00 Seconds
1508.451 Million cell updates/sec

Tabular output not generated.

Title: >US-08-908-884-13
Description: (1-2172) from US08908884.seq
Perfect Score: 2172
N.A. Sequence: 1 GTGACTTTCTAACTATGGCT.....ATTTGAAAAAAA 2172
Comp: CACTGAAAGATTGATAACCGA.....TAAACTTTTTTTTTTTT

Scoring table: TABLE default
Gap 6

Nmatch STD : Dbase 0; Query 0

Searched: 532259 seqs, 918475165 bases x 2

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: embl155
1:em_ba 2:em_htg 3:em_hum1 4:em_hum2 5:em_in 6:em_om
7:em_or 8:em_ov 9:em_pat 10:em_pl 11:em_ro

Database: genbank107
12:gb_ba 13:gb_htg 14:gb_in 15:gb_om 16:gb_ov 17:gb_pat
18:gb_ph 19:gb_pl 20:gb_pr1 21:gb_pr2 22:gb_ro 23:gb_st
24:gb_sts 25:gb_sy 26:gb_un 27:gb_vi

Statistics: Mean 11.830; Variance 6.147; scale 1.925

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result	No.	Score	Query		DB	ID	Description	Pred.	No.
			%	Match					
	1	328	15.1	2104	19	ATU76707	Arabidopsis thaliana r	1.27e-228	
	2	165	7.6	5655	19	ATU87794	Arabidopsis thaliana t	7.33e-99	
c	3	74	3.4	7218	17	I66494	Sequence 14 from paten	2.44e-30	
c	4	68	3.1	7218	17	I66494	Sequence 14 from paten	4.07e-26	
c	5	40	1.8	10772	14	AF012089	Drosophila melanogaste	1.15e-07	
c	6	37	1.7	10772	14	AF012089	Drosophila melanogaste	6.99e-06	
c	7	34	1.6	354	14	OFU89259	Oxytricha fallax 57kD	3.70e-04	
c	8	33	1.5	215	17	I28278	Sequence 5 from patent	1.34e-03	
c	9	33	1.5	103548	13	HS292E10	Human DNA sequence ***	1.34e-03	
c	10	32	1.5	152580	13	HS366L4	Human DNA sequence ***	4.79e-03	
c	11	28	1.3	375	24	G25545	human STS EST118797.	6.22e-01	
c	12	28	1.3	529	24	G27780	human STS SHGC-33252.	6.22e-01	
c	13	28	1.3	125536	13	AC003021	*** SEQUENCING IN PROG	6.22e-01	
c	14	28	1.3	168133	13	HS500L14	Human DNA sequence ***	6.22e-01	
c	15	28	1.3	219090	13	HS212A2	Human DNA sequence ***	6.22e-01	
c	16	26	1.2	30	9	A62994	Sequence 6 from Patent	6.11e+00	
c	17	26	1.2	60	9	A62989	Sequence 1 from Patent	6.11e+00	
c	18	25	1.2	111	9	A60119	Sequence 7 from Patent	1.83e+01	
c	19	26	1.2	215	17	I28278	Sequence 5 from patent	6.11e+00	
c	20	25	1.2	267	24	BLYBD	Hordeum vulgare (clone	1.83e+01	
c	21	25	1.2	400	24	G10880	human STS SHGC-9303 cl	1.83e+01	
c	22	26	1.2	565	17	E04076	gDNA encoding envelope	6.11e+00	
c	23	27	1.2	602	24	HUMUT6251	Human STS UT6251.	1.98e+00	
c	24	26	1.2	818	19	BOU93838	Blechnum occidentale A	6.11e+00	
c	25	25	1.2	1122	14	AF020286	Dictyostelium discoide	1.83e+01	
c	26	26	1.2	1416	19	LUU59284	Linum usitatissimum pe	6.11e+00	
c	27	25	1.2	1577	19	DDINDKM	Dictyostelium discoide	1.83e+01	
c	28	26	1.2	1918	14	S72639	PspB=prespore-specific	6.11e+00	
c	29	25	1.2	2222	14	DDSP96A	Dictyostelium discoide	1.83e+01	
c	30	25	1.2	2493	20	HUMIGHAD	Human Ig rearranged H-	1.83e+01	
c	31	25	1.2	2904	21	HSAC001456	Homo sapiens (subclone	1.83e+01	
c	32	25	1.2	8042	21	HSHCFC1	H.sapiens HCFC1 gene.	1.83e+01	
c	33	26	1.2	9705	19	CHLC27	C.ellipsoidea rbcL, rp	6.11e+00	
c	34	25	1.2	27119	14	CEF08B12	Caenorhabditis elegans	1.83e+01	
c	35	25	1.2	40977	21	HSAC000362	Human cosmid g1346a094	1.83e+01	
c	36	25	1.2	41702	21	HSAC000366	Human cosmid g1862x046	1.83e+01	
c	37	26	1.2	78794	21	AC003990	Human BAC clone RG035M	6.11e+00	
c	38	25	1.2	116019	21	AC004629	Homo sapiens chromosom	1.83e+01	
c	39	26	1.2	120007	21	AF064864	Homo sapiens chromosom	6.11e+00	
c	40	26	1.2	133469	13	HSAC000383	*** SEQUENCING IN PROG	6.11e+00	
c	41	25	1.2	145717	13	HS108K11	Human DNA sequence ***	1.83e+01	
c	42	26	1.2	150613	19	AB001684	Chlorella vulgaris C-2	6.11e+00	
c	43	25	1.2	156601	21	HS212G6	Homo sapiens DNA seque	1.83e+01	
c	44	25	1.2	160000	13	AC004062	*** SEQUENCING IN PROG	1.83e+01	
c	45	25	1.2	224036	13	AC004787	*** SEQUENCING IN PROG	1.83e+01	

ALIGNMENTS

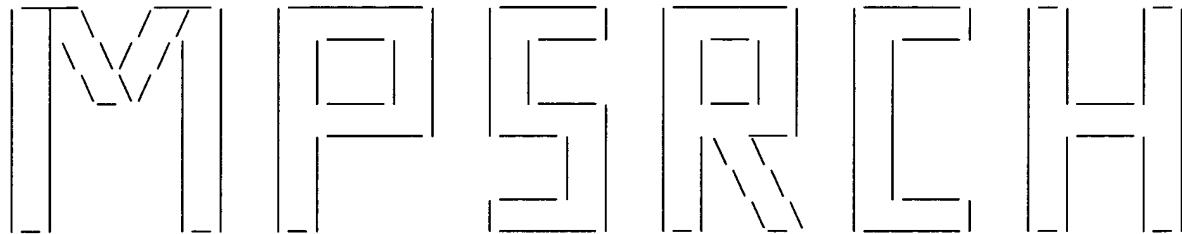
LOCUS ATU76707 2104 bp mRNA PLN 14-JAN-1997
 DEFINITION Arabidopsis thaliana regulatory protein NPR1 (NPR1) mRNA, complete cds.
 ACCESSION U76707
 NID g1773294
 KEYWORDS .
 SOURCE thale cress.
 ORGANISM Arabidopsis thaliana
 Eukaryotae; mitochondrial eukaryotes; Viridiplantae;
 Charophyta/Embryophyta group; Embryophyta; Magnoliophyta;
 Magnoliopsida; Capparales; Brassicaceae; Arabidopsis.
 REFERENCE 1 (bases 1 to 2104)
 AUTHORS Cao,H., Glazebrook,J., Clarke,J.D., Volko,S. and Dong,X.
 TITLE The Arabidopsis NPR1 gene that controls systemic acquired resistance encodes a novel protein containing ankyrin repeats
 JOURNAL Cell 88 (1), 57-63 (1997)
 MEDLINE 97148688
 REFERENCE 2 (bases 1 to 2104)
 AUTHORS Cao,H., Glazebrook,J., Clarke,J.D., Volko,S. and Dong,X.
 TITLE Direct Submission
 JOURNAL Submitted (30-OCT-1996) DCMB Group, Botany, Duke University, LSRC Building, Research Dr., Durham, NC 27708-1000, USA
 FEATURES Location/Qualifiers
 source 1. .2104
 /organism="Arabidopsis thaliana"
 /strain="ecotype Columbia"
 /db_xref="taxon:3702"
 /chromosome="I"
 /map="between GAP-B and m315"
 gene 1. .2104
 /gene="NPR1"
 CDS 93. .1874
 /gene="NPR1"
 /note="ankyrin repeat-containing protein; gene controlling systemic acquired resistance; SAR"
 /codon_start=1
 /product="regulatory protein NPR1"
 /db_xref="PID:g1773295"
 /translation="MDTTIDGFADSYEISSTSFVATDNTDSSIVYLAEEQVLTGPDV
 ALQLLSNSFESVFDSPDDFYSDAKLVLSDGREVSFHRCVLSARSSFFKSALAAKKEK
 DSNNTAAVKLELKEIAKKDYEVGFDSSVVTLAYVYSSRVRPPKGVSECADENCHVAC
 RPAVDFMLEVLYLAFIFKIPELITLYQRHLLDVVDKVVIEDTLVILKLANICGKACMK
 LLDRCKEIIVKSNVDMVSLEKSLPEELVKEIIDRRKELGLEVPKVKKHVSNVHKALDS
 DDIELVKLLLKEDHTNLDDACALHFAYCNVKTATDLLKLDLADVNHRNPRGYTVLH
 VAAMRKEPQLLISLLEKGASASEATLEGRTALMIAKQATMAVECNNIPEQCKHSLKGR
 LCVEILEQEDKREQIPRDVPPSFAVAADELKMTLLDLENRVALAQRILFPTEAQAAAMEI
 AEMKGTCFIVTSLEPDRLTGTKRTSPGVKIAPFRILEEHQSRLKALSKTVELGKRFF
 PRCSAVLDQIMNCEDLTQLACGEDDTAEKRLQKKQRYMEIQETLKKAFSEDNLELGNS
 SLSTDSTSSTSKSTGGKRSNRKLSHRRR"
 BASE COUNT 606 a 431 c 482 g 585 t
 ORIGIN

Query Match 15.1%; Score 328; DB 19; Length 2104;
 Best Local Similarity 64.8%; Pred. No. 1.27e-228;

Matches 882; Conservative 0; Mismatches 464; Indels 15; Gaps 5;

Db 423 AAGCTCGAGCTTAAGGAGATTGCCAAGGATTACGAAGTCGGTTCGATTGGTGTGACT 482
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 555 AAGGTGGAATTGAAGGAGGTGATGAAAGAGCATGAGGTGAGCTATGCTGTAATGAGT 614
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 483 GTTTGGCTTATGTTACAGCAGCAGAGTGAGACCGCCGCCTAAAGGAGTTCTGAATGC 542
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 615 GTATTGGCTTATTGTATAGTGGTAAAGTTAGGCCTTCACCTAAAGATGTGTGTGTTGT 674
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 543 GCAGACGAGAATTGCTGCCACGTGGCTGCCGGCGCGGTGGATTTCATGTTGGAGGTT 602
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 675 GTGGACAATGACTGCTCTCATGTGGCTGTAGGCCAGCTGTGGCATTCTGGTTGAGGTT 734
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 603 CTCTATTTGGCTTCATCTCAAGATCCCTGAATTAAATTACTCTCTATCAGAGGCACCTA 662
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 735 TTGTACACATCATTACCTTCAGATCTCTGAATTGGTTGACAAGTTCAAGAGACACCTA 794
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 663 TTGGACGTTTAGACAAAGTTAGAGGACACATTGGTTATACTCAAGCTTGCTAAT 722
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 795 CTGGATATTCTTGACAAAAGTGCAGCAGACGATGTAATGATGGTTTATCTGTTGCAAAC 854
||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 723 ATATGTGGTAAAGCTTGTATGAAGCTATTGGATAGATGAAAGAGATTATTGTCAAGTCT 782
||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 855 ATTTGTGGTAAAGCATGCGAGAGATTGCTTCAAGCTGCATTGAGATTATTGTCAAGTCT 914
||| ||| ||| ||| ||| ||| ||| ||| |||
Db 783 AATGTAGATATGGTTAGTCTTGAAGCTATTGGCTTAAAGAGATAATT 842
||| ||| ||| ||| ||| ||| ||| |||
Qy 915 AATGTTGATATCATAACCCTTGATAAAGCCTTGCTCATGACATTGAAACAAATTACT 974
||| ||| ||| ||| ||| ||| |||
Db 843 GATAGACGTAAGAGCTTGGTTGGAGGTACCT-AAAGTAAAG-----AAACAT 890
||| ||| ||| ||| ||| ||| ||| |||
Qy 975 GATTACGAGCGGAACCTGGCTACAAGGGCCTGAAAGCAACGGTTTCCTGATAAACAT 1034
||| ||| ||| ||| ||| ||| |||
Db 891 GTCTCGAATGTACATAAGGCACCTGACTCGGATGATATTGAGTTAGTCAGTTGCTTTG 950
||| ||| ||| ||| ||| ||| ||| |||
Qy 1035 GTTAAGAGGATACATAGGGCATTGGATTCTGATGATGTTGAATTACTACAAATGTTGCTA 1094
||| ||| ||| ||| ||| ||| |||
Db 951 AAAGAGGATCACACCAATCTAGATGATGCGTGTGCTCTCATTGCTGTTGCATATTGC 1010
||| ||| ||| ||| ||| ||| ||| |||
Qy 1095 AGAGAGGGCATACTACCCTAGATGATGCATATGCTCTCATTATGCTGTAGCGTATTGC 1154
||| ||| ||| ||| ||| ||| |||
Db 1011 AATGTGAAGACCGAACAGATCTTAAACTGATCTGCCGATGTCACCCATAGGAAT 1070
||| ||| ||| ||| ||| ||| ||| |||
Qy 1155 GATGCAAAGACTACAGCAGAACTCTAGATCTGCACTGCTGATATTAATCATCAAAT 1214
||| ||| ||| ||| ||| ||| |||
Db 1071 CCGAGGGATACGGTGCCTCATGTTGCTGCGATGCGGAAGGAGCCACAATTGATACTA 1130
||| ||| ||| ||| ||| ||| ||| |||
Qy 1215 TCAAGGGATACACGGTGCATGTTGCAGCCATGAGGAAAGAGCCTAAATTGTAGTG 1274
||| ||| ||| ||| ||| ||| |||
Db 1131 TCTCTATTGGAAAAAGGTGCAAGTGCATCAGAACCAACTTGGAAAGGTAGAACCGCACTC 1190
||| ||| ||| ||| ||| ||| |||
Qy 1275 TCCCTTTAACCAAAGGAGCTAGACCTCTGATCTGACATCCGATGGAAGAAAAGCACTT 1334

Db 1191 ATGATCGAAAACAAGCCACTATGGCGGTTGAATGTAATAATATCCCGAGCAATGCAAG 1250
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1335 CAAATGCCAAGAGGCTCACTAGGCTTGTGGATTCAAGTCTCCGGAGGAAGGAAAA 1394
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1251 CATTCTCTCAAAGGCCACTATGTGTAGAAATACTAGAGCAAGAACGAGAACAA 1310
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1395 TCTGCTTCGAATGATCGGTTATGCATTGAGATTCTGGAGCAAGCAGAAAGAAGACCT 1454
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1311 ATTCCCTAGAGATGTTCCCTCCCTTTGCAGTGGCGGCCATGAATTGAAGATGACGCTG 1370
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1455 CTGCTAGGAGAACGTTCTGTATCTCTGCTATGGCAGGCGATGATTGCGTATGAAGCTG 1514
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1371 CTCGATCTTGAAAATAGAGTTGCACTTGCTAACGTCTTTCCAACGGAAGCACAGCT 1430
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1515 TTATACCTTGAAAATAGAGTTGGCCTGGCTAAACTCCTTTCCAATGGAAGCTAAAGTT 1574
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1431 GCAATGGAGATCGCCGAAATGAAGGAAACATGTGAGTTCATAGTGACTIONCCTCGAGCCT 1490
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1575 GCAATGGACATTGCTCAAGTTGATGGCACTTCTGAGTTCCCCTGGCTAGCATCG-GCAA 1633
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1491 GACCGTCTCACTGGTACGAAGAGAACATCACCGGGTGTAAAGATAGCACCTTCAGAAC 1550
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1634 AAA-GA-TGGCTAATGCACAGAGGACAACAGTAGATTGAAACGAGGCTCCTTCAAGATA 1691
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1551 CTAGAAGAGCATCAAAGTAGACTAAAAGCGCTTCTAAAACCGTGGAACTCGGGAAACGA 1610
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1692 AAAGAGGAGCACTTGAATCGGCTTAGAGCACTCTCTAGAACTGTAGAACTTGGAAAACGC 1751
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1611 TTCTTCCCGCGCTGTTGGCAGTGCTCGACCAGATTATGAACTGTGAGGACTTGACTCAA 1670
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1752 TTCTTCCACGTTGTTCAGAAGTTCTAAATAAGATCATGGATGCTGATGACTTGTCTGAG 1811
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1671 CTGGCTTGCAGAGAACGACACTGCTGAGAACGACTACAAAAGAACGAAAGGTACATG 1730
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1812 ATAGCTTACATGGGAATGATAACGGCAGAACAGCGTCAACTGAAGAACGAAAGGTACATG 1871
||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1731 GAAATACAAGAGACACTAAAGAACGGCTTGTGAGGACAA 1771
||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1872 GAACCTCAAGAAATTCTGACTAAAGCATTCACTGAGGATAA 1912



(TM)

Release 3.1A John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

MPsrch_nn n.a. - n.a. database search, using Smith-Waterman algorithm
Run on: Mon Oct 5 22:08:16 1998; MasPar time 313.34 Seconds
943.086 Million cell updates/sec
Tabular output not generated.

Title: >US-08-908-884-13
Description: (1-2172) from US08908884.seq
Perfect Score: 2172
N.A. Sequence: 1 GTGACTTTCTAACTATGGCT.....ATTTGAAAAAAAAAAAAAAA 2172
Comp: CACTGAAAGATTGATAACCGA.....TAAACTTTTTTTTTTTTTTTTT

Scoring table: TABLE default
Gap 6

Nmatch STD : Dbase 0; Query 0

Searched: 188442 seqs, 68026449 bases x 2

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: n-geneseq32
1:part1 2:part2 3:part3 4:part4 5:part5 6:part6 7:part7
8:part8 9:part9 10:part10 11:part11 12:part12 13:part13
14:part14 15:part15 16:part16 17:part17 18:part18
19:part19 20:part20 21:part21 22:part22 23:part23
24:part24 25:part25 26:part26 27:part27 28:part28
29:part29 30:part30 31:part31 32:part32 33:part33
34:part34 35:part35 36:part36 37:part37 38:part38
39:part39 40:part40

Statistics: Mean 9.671; Variance 6.087; scale 1.589

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result	No.	Score	Query				Description	Pred.	No.
			Match	Length	DB	ID			
	1	165	7.6	5655	40	V04632	Arabidopsis thaliana	1.35e-85	
c	2	165	7.6	9919	40	V04631	Arabidopsis thaliana	1.35e-85	
c	3	43	2.0	204	1	N81164	Base substituted E.co	9.86e-10	
	4	40	1.8	91	9	Q51746	Oligonucleotide probe	4.18e-08	
	5	39	1.8	204	1	N81164	Base substituted E.co	1.43e-07	
c	6	38	1.7	91	9	Q51746	Oligonucleotide probe	4.87e-07	
	7	35	1.6	114	12	Q70467	Generic DNA sequence	1.80e-05	
	8	34	1.6	114	12	Q70469	Generic DNA sequence	5.87e-05	
	9	34	1.6	114	12	Q70465	Generic DNA sequence	5.87e-05	
c	10	35	1.6	114	12	Q70465	Generic DNA sequence	1.80e-05	
c	11	34	1.6	114	12	Q70468	Generic DNA sequence	5.87e-05	

C	12	34	1.6	114	12	Q70466	Generic DNA sequence	5.87e-05
	13	32	1.5	114	12	Q70468	Generic DNA sequence	6.02e-04
C	14	32	1.5	114	12	Q70467	Generic DNA sequence	6.02e-04
	15	32	1.5	501	3	N50024	Sequence encoding new	6.02e-04
	16	32	1.5	501	3	N50026	Sequence encoding new	6.02e-04
	17	30	1.4	114	12	Q70466	Generic DNA sequence	5.86e-03
	18	30	1.4	114	12	Q70470	Generic DNA sequence	5.86e-03
C	19	31	1.4	114	12	Q70469	Generic DNA sequence	1.89e-03
	20	31	1.4	3871	2	N71302	HSV-1 gB and surround	1.89e-03
	21	28	1.3	39	7	Q51787	Mixed oligonucleotide	5.37e-02
C	22	29	1.3	114	12	Q70470	Generic DNA sequence	1.79e-02
C	23	28	1.3	114	12	Q70472	Generic DNA sequence	5.37e-02
	24	28	1.3	498	3	N50034	Sequence encoding new	5.37e-02
	25	29	1.3	501	3	N50030	Sequence encoding new	1.79e-02
	26	28	1.3	501	3	N50029	Sequence encoding new	5.37e-02
	27	28	1.3	501	3	N50031	Sequence encoding new	5.37e-02
	28	28	1.3	501	3	N50023	Sequence encoding new	5.37e-02
	29	28	1.3	501	3	N50028	Sequence encoding new	5.37e-02
	30	28	1.3	501	3	N50032	Sequence encoding new	5.37e-02
C	31	25	1.2	75	21	T13612	DC43 TSAR library gen	1.31e+00
	32	25	1.2	111	36	T96515	MSRV-1B general conse	1.31e+00
	33	26	1.2	114	12	Q70471	Generic DNA sequence	4.60e-01
C	34	26	1.2	114	12	Q70473	Generic DNA sequence	4.60e-01
	35	26	1.2	140	32	T76368	Human IL-8 receptor-a	4.60e-01
	36	26	1.2	172	32	T76363	Human interleukin 8 a	4.60e-01
	37	27	1.2	178	32	T76405	Human endothelin-1 an	1.59e-01
C	38	25	1.2	178	32	T76405	Human endothelin-1 an	1.31e+00
	39	27	1.2	501	3	N50025	Sequence encoding new	1.59e-01
	40	26	1.2	501	3	N50027	Sequence encoding new	4.60e-01
	41	26	1.2	501	3	N50033	Sequence encoding new	4.60e-01
C	42	25	1.2	501	3	N50026	Sequence encoding new	1.31e+00
C	43	25	1.2	501	3	N50030	Sequence encoding new	1.31e+00
C	44	25	1.2	501	3	N50032	Sequence encoding new	1.31e+00
C	45	25	1.2	501	3	N50025	Sequence encoding new	1.31e+00

ALIGNMENTS

RESULT 1

ID V04632 standard; DNA; 5655 BP.

AC V04632;

DT 17-JUN-1998 (first entry)

DE Arabidopsis thaliana NIM1 gene.

KW NIM1; noninducible immunity; disease resistance; plants;

KW SAR gene expression; ss.

OS Arabidopsis thaliana.

FH Key Location/Qualifiers

FT CDS 2787..4866)

FT /*tag= a

FT /product= "NIM1 protein"

FT exon 2787..3347

FT /*tag= b

FT /number= 1

FT exon 3427..4162

Db 3838 aggaatccgaggggatatacggtgcttcattgttgcgtgcgtgcggaaaggagccacaattg 3897
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1209 CAAAATTCAAGGGATACACGGTGCTGCATGTTGCAGCCATGAGGAAAGAGCCTAAAATT 1268

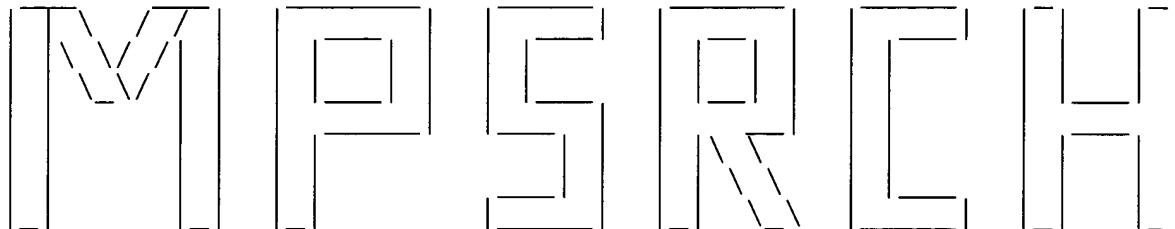
Db 3898 atactatctctattggaaaaagggtgcagaatgcatacagaaggcaacttttggaaaggtagaacc 3957
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1269 GTAGTGTCCCTTTAACCAAAGGAGCTAGACCTCTGATCTGACATCCGATGGAAGAAAA 1328

Db 3958 gcactcatgatcgcaaaaacaaggccactatggcggttgaatgtataataatatcccggagcaa 4017
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1329 GCACTTCAAATGCCAAGAGGCTCACTAGGCTTGTGGATTTCAGTAAGTCTCCGGAGGAA 1388

Db 4018 tgcaaggcattctctcaaaggccgactatgtgttagaaatactagagcaagaagacaaacga 4077
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1389 GGAAAATCTGCTTCGAATGATCGGTTATGCATTGAGATTCTGGAGCAAGCAGAAAGAAGA 1448

Db 4078 gaacaaattcttagagatgttccctctttgcagtggcgccgatgaattgaagatg 4137
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1449 GACCCTCTGCTAGGAGAACGCTCTGTATCTCTGCTATGGCAGGCGATGATTGCGTATG 1508

Db 4138 acgctgctcgatcttggaaaatagag 4162
||| ||| ||| ||| ||| |||
Qy 1509 AAGCTGTTATAACCTTGAAAATAGAG 1533



(TM)

Release 3.1A John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

MPsrch_pp protein - protein database search, using Smith-Waterman algorithm

Run on: Mon Oct 5 10:54:52 1998; MasPar time 28.23 Seconds
877.191 Million cell updates/sec

Tabular output not generated.

Title: >US-08-908-884-14

Description: (1-588) from US08908884.pep

Perfect Score: 4115

Sequence: 1 MDNSRTAFSDSNDISGSSSI.....SSCSSTSKGVDKPNKLPFRK 588

Scoring table: PAM 150

Gap 11

Searched: 140555 seqs, 42109429 residues

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database: sptrembl6

1:sp_fungi 2:sp_human 3:sp_invertebrate 4:sp_mammal
5:sp_mhc 6:sp_organelle 7:sp_phage 8:sp_plant
9:sp_bacteria 10:sp_rat 11:sp_virus 12:sp_vertebrate
13:sp_unclassified

Statistics: Mean 50.515; Variance 104.765; scale 0.482

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query		Length	DB	ID	Description	Pred. No.
		%	Match					
1	2184	53.1	593	8	P93002		REGULATORY PROTEIN NPR	0.00e+00
2	185	4.5	1088	2	Q13484		ANKYRIN G119.	1.55e-11
3	184	4.5	1765	10	Q61308		ANKYRIN 3.	2.21e-11
4	184	4.5	1940	10	Q61309		ANKYRIN 3.	2.21e-11
5	184	4.5	1943	10	Q61307		ANKYRIN 3, EPITHELIAL.	2.21e-11
6	184	4.5	1961	10	Q61310		ANKYRIN 3.	2.21e-11
7	185	4.5	4377	2	Q12955		ANKYRIN G.	1.55e-11
8	172	4.2	1214	3	Q25338		DELTA-LATROINSECTOTOXI	1.53e-09
9	173	4.2	1719	2	Q13768		ALT. ANKYRIN (VARIANT	1.08e-09
10	173	4.2	1856	2	Q99407		ANKYRIN.	1.08e-09
11	169	4.1	843	10	P97582		ANKYRIN (FRAGMENT).	4.33e-09
12	170	4.1	1098	10	Q61304		ERYTHROID ANKYRIN (FRA	3.06e-09
13	170	4.1	1848	10	Q61302		ERYTHROID ANKYRIN.	3.06e-09
14	147	3.6	1151	1	O13987		HYPOTHETICAL 127.7 KD	7.15e-06
15	149	3.6	1411	3	Q02989		ALPHA-LATROINSECTOTOXI	3.71e-06
16	145	3.5	1549	3	Q24241		ANKYRIN.	1.37e-05
17	139	3.4	813	3	Q24313		DNA BINDING PROTEIN.	9.41e-05
18	134	3.3	448	12	Q92010		ZINC FINGER 5 PROTEIN	4.54e-04
19	135	3.3	813	3	P90538		MIGA.	3.32e-04
20	133	3.2	449	2	O00403		ZINC FINGER 5 PROTEIN.	6.20e-04
21	133	3.2	449	10	Q08376		ZINC FINGER PROTEIN ZF	6.20e-04
22	131	3.2	1786	3	Q17344		UNC-44 (FRAGMENT).	1.15e-03
23	131	3.2	1809	3	Q17487		E. ELEGANS ANKYRIN-REL	1.15e-03
24	131	3.2	1815	3	Q17488		C. ELEGANS ANKYRIN-REL	1.15e-03
25	131	3.2	1867	3	Q17486		C. ELEGANS ANKYRIN-REL	1.15e-03
26	131	3.2	2039	3	Q17489		C. ELEGANS ANKYRIN-REL	1.15e-03
27	131	3.2	6994	3	Q17343		UNC-44 ANKYRINS.	1.15e-03
28	127	3.1	211	11	O41154		A672R PROTEIN.	3.90e-03
29	129	3.1	611	3	Q18670		HYPOTHETICAL PROTEIN C	2.12e-03
30	128	3.1	882	2	O15084		KIAA0379 (FRAGMENT).	2.88e-03
31	123	3.0	341	9	O06637		PUTATIVE FLAGELLA-RELA	1.30e-02
32	125	3.0	581	3	O16313		C05C8.6 PROTEIN.	7.13e-03
33	121	2.9	151	2	Q92527		ANKYRIN MOTIF.	2.34e-02
34	118	2.9	520	9	Q47163		(CTR5X).	5.62e-02
35	120	2.9	552	2	Q14776		LZTR-1.	3.14e-02
36	121	2.9	625	3	O18147		T27E9.4.	2.34e-02
37	120	2.9	1640	3	O17112		F39E9.2 PROTEIN.	3.14e-02
38	115	2.8	186	9	O24706		ADENYLATE KINASE.	1.33e-01
39	116	2.8	251	10	Q62422		SH3P2 (FRAGMENT).	1.00e-01
40	116	2.8	443	9	Q54703		EMM18.1.	1.00e-01
41	116	2.8	508	3	Q20604		F49E11.1.	1.00e-01
42	116	2.8	1005	12	O13075		NUCLEAR FACTOR NF-KB1.	1.00e-01
43	116	2.8	1401	3	Q25328		ALPHA-LATROTOXIN PRECU	1.00e-01
44	113	2.7	474	11	P87621		41KBP FRAGMENT FROM LE	2.34e-01
45	113	2.7	1001	2	Q14349		G9A PROTEIN CONTAINING	2.34e-01

ALIGNMENTS

RESULT 1

ID P93002 PRELIMINARY; PRT; 593 AA.
AC P93002; 004742;
DT 01-MAY-1997 (TREMBLREL. 03, CREATED)
DT 01-MAY-1997 (TREMBLREL. 03, LAST SEQUENCE UPDATE)
DT 01-JAN-1998 (TREMBLREL. 05, LAST ANNOTATION UPDATE)
DE REGULATORY PROTEIN NPR1.
GN NPR1 OR NIM1.
OS ARABIDOPSIS THALIANA (MOUSE-EAR CRESS).
OC EUKARYOTA; PLANTA; EMBRYOPHYTA; ANGIOSPERMAE; DICOTYLEDONEAE;
OC CAPPARALES; CRUCIFERAE.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ECOTYPE COLUMBIA;
RX MEDLINE; 97148688.
RA CAO H., GLAZEBROOK J., CLARKE J.D., VOLKO S., DONG X.;
RL CELL 88:57-63 (1997).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. COLUMBIA;
RA RYALS J., WEYMANN K., LAWTON K., FRIEDRICH L., ELLIS D., STEINER H.Y.,
RA JOHNSON J., DELANEY T.P., JESSE T., VOS P., UKNES S.;
RL PLANT CELL 9:0-0(0).
DR EMBL; U76707; G1773295; -.
DR EMBL; U87794; G1916912; -.
SQ SEQUENCE 593 AA; 66031 MW; 1BE1666B CRC32;

Query Match 53.1%; Score 2184; DB 8; Length 593;
Best Local Similarity 51.2%; Pred. No. 0.00e+00;
Matches 307; Conservative 150; Mismatches 123; Indels 20; Gaps 14;

Db 1 MDTTIDGFADSYEISSTSFVATDNTDSSIVYLAEEQVLTGPDVSALQLLSNSFESVFD-S 59
|| : : | : || : | : | : | : : | : : : | : : : : | | : : | | : | | |
Qy 1 MDNSRTAFSDSNDISGSSSIICCIGGGMTE-FFSPE-T-SPAEITSLKRLSETLESIFDAS 57

Db 60 -PD-DFYSDAKLVLSDG-REVSFHRCVLSARSSFFKSALAAKKEKDSNNTAAVKLELKE 116
| : | :: : | | | | : | : | : | | | | : | : | | | | : | : | | | |
Qy 58 LPEFDYFADAKLVVSGPCKEIPVHRCILSARSPFFKN-LFCGKKEKNS---S-KVELKE 111

Db 117 IAKDYEVGFDSVVTVLAYVYSSRVRRPPPCKVSECADENCCHVACRPAVDFMLEVLYLAFI 176
: | : | | : | : | | | : | | | | | | | | | | | | | | | | | | | | | | |
Qy 112 VMKEHEVSYDAVMSVLAYLYSGKVRPSPKDVCVCVDNDCSHVACRPAVAFLVEVLYTSFT 171

Db 177 FKIPELITLYQRHLLDVVKVIEDTLVILKLANICGKACMKLLDRCKEIIVKSNVDMVS 236
| | : | : | : | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 172 FQISELVDKFQRHLLDILDKTAADDVMMVLSVANICGKACERLLSSCIEIIIVKSNVDIIT 231

Db 237 LEKSLPEELVKEIIDRRKELGLEVPKVK---KHVSNVHKALDSDDIELVKLLLKEDHTN 292
| : | : | : | : | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 232 LDKALPHDIVKQITDSRAELGLQGPESNGFPDKHVKRIHRALDSDDVELLQMLLREGHTT 291

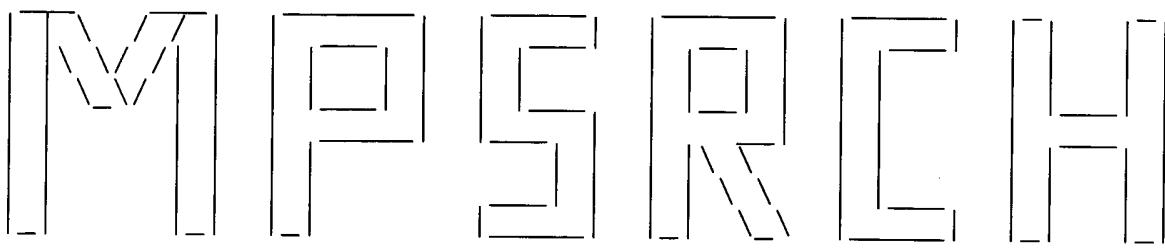
Db 293 LDDACALHFAVAYCNVKTATDLLKLDLADVNRHRNPRGYTVLHVAAMRKEPQLILSLLKG 352
| | | | | : | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 292 LDDAYALHYAVAYCDAKTTAELLDLALADINHQNSRGYTVLHVAAMRKEPKIVVSLLTKG 351

Db 353 ASASEATLEGR TALMIAKQATMAVECNNIPEQCKHSLKGRLCVEILEQEDKREQIPRDVP 412
| : | : | : || | : | | : | : : | | : | : | : | : | : :
Qy 352 ARPSDLTSDGRKALQIAKRLTRLVDFSKSPEEGKSASNDRLCIEILEQAERRDPLLGEAS 411

Db 413 PSFAVAADELKMTLLDLENRVALAQRLFPTEAQAA MEIAEMKGTCEFIVTSLEPDRLTGT 472
| : | : | : | : | | | | : | | | | | | | : | | | | : | : : :
Qy 412 VSLAMAGDDL RRMKLLYLENRVGLAKLLFPMEAKVAMDIAQVDGTSEFPLASIG-KKMANA 470

Db 473 KRTSPGVKIA PFRILEEHQSRLKALSKTVELGKRFFPRCSAVLDQIMNCEDLTQLACGED 532
| | : | : | | | : | | | : | | | | | | | | | | : | | : | | : | :
Qy 471 QRTTVDLNEAPFKIKEEHLNRLRAL SRTVELGKRFFPRCSEVLNKIMDADDLSEIAYMGN 530

Db 533 DTAEKRLQKKQRYMEIQETLKKA FSEDNLELGNSSLTDSTSSTSKSTGGKRSNRKLSHRR 592
| | | | | | | | : | | | | | : | : | : | | | | | | : | : | : | :
Qy 531 DTAEERQLKKQRYMELQEILTKAFTEDKEEYDKTN-NI SSSCSSTS KGVDKPN-KLPFRK 588



(TM)

Release 3.1A John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

MPsrch_pp protein - protein database search, using Smith-Waterman algorithm

Run on: Mon Oct 5 10:59:59 1998; MasPar time 16.09 Seconds
591.174 Million cell updates/sec

Tabular output not generated.

Title: >US-08-908-884-14
Description: (1-588) from US08908884.pep
Perfect Score: 4115
Sequence: 1 MDNSRTAFSDSNDISGSSSI.....SSCSSTSKGVDKPNKLPFRK 588

Scoring table: PAM 150
Gap 11

Searched: 131922 seqs, 16180660 residues

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: a-geneseq32
1:part1 2:part2 3:part3 4:part4 5:part5 6:part6 7:part7
8:part8 9:part9 10:part10 11:part11 12:part12 13:part13
14:part14 15:part15 16:part16 17:part17 18:part18
19:part19 20:part20 21:part21 22:part22 23:part23
24:part24 25:part25 26:part26 27:part27 28:part28
29:part29

Statistics: Mean 36.213; Variance 178.711; scale 0.203

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Query	Description	Pred. No.
						%		

1	2184	53.1	593	29	W23963	Arabidopsis thaliana	3.65e-174
2	172	4.2	991	16	R80096	Black widow spider de	5.09e-04
3	172	4.2	1214	16	R80097	Black widow spider de	5.09e-04
4	134	3.3	553	23	W15103	Myxoma virus MA55 gen	2.45e-01
5	117	2.8	226	21	W15483	Human P28.	3.38e+00
6	116	2.8	251	25	W05386	Mouse SH3P2 protein.	3.93e+00
7	112	2.7	943	10	R51499	Cyanobacteria replica	7.17e+00
8	105	2.6	231	9	R43385	Product of homeotic g	2.01e+01
9	105	2.6	394	22	W17845	Cytosolic phospholipa	2.01e+01
10	105	2.6	687	22	W17847	Cytosolic phospholipa	2.01e+01
11	105	2.6	688	22	W17848	Cytosolic phospholipa	2.01e+01
12	106	2.6	969	7	R33213	Sequence of the p105	1.74e+01
13	105	2.6	1078	6	R28963	Notch hN3k full lengt	2.01e+01
14	104	2.5	589	16	R94389	Mouse neural cell pro	2.33e+01
15	104	2.5	589	16	R94386	Human neural cell pro	2.33e+01
16	103	2.5	718	22	W01537	Bovine ribonuclease L	2.69e+01
17	104	2.5	752	17	R83018	Calcium-independent c	2.33e+01
18	104	2.5	752	20	W01479	Calcium-independent c	2.33e+01
19	104	2.5	752	22	W17849	Hamster cytosolic pho	2.33e+01
20	104	2.5	752	21	W13163	Ca-independent phosph	2.33e+01
21	98	2.4	229	25	W05403	Human clone 5 protein	5.51e+01
22	97	2.4	318	20	W06318	Human mitogen-activat	6.34e+01
23	97	2.4	657	6	R28964	Notch hN5k full lengt	6.34e+01
24	98	2.4	955	8	R42249	Human p50 protein DNA	5.51e+01
25	98	2.4	955	8	R42237	Human p50 protein DNA	5.51e+01
26	98	2.4	955	8	R42253	Human p50 protein DNA	5.51e+01
27	98	2.4	955	8	R42251	Human p50 protein DNA	5.51e+01
28	98	2.4	955	8	R42236	Human p50 protein DNA	5.51e+01
29	98	2.4	955	8	R42252	Human p50 protein DNA	5.51e+01
30	98	2.4	955	8	R42239	Human p50 protein DNA	5.51e+01
31	98	2.4	955	8	R42246	Human p50 protein DNA	5.51e+01
32	98	2.4	955	8	R42234	Human p50 protein DNA	5.51e+01
33	98	2.4	955	8	R42238	Human p50 protein DNA	5.51e+01
34	98	2.4	955	8	R42241	Human p50 protein DNA	5.51e+01
35	98	2.4	955	8	R42242	Human p50 protein DNA	5.51e+01
36	98	2.4	955	8	R42088	Human p50 or KBF1 pro	5.51e+01
37	98	2.4	955	8	R42250	Human p50 protein DNA	5.51e+01
38	98	2.4	955	8	R42240	Human p50 protein DNA	5.51e+01
39	98	2.4	955	8	R42245	Human p50 protein DNA	5.51e+01
40	98	2.4	955	8	R42247	Human p50 protein DNA	5.51e+01
41	98	2.4	955	8	R42244	Human p50 protein DNA	5.51e+01
42	98	2.4	955	8	R42243	Human p50 protein DNA	5.51e+01
43	98	2.4	955	8	R42235	Human p50 protein DNA	5.51e+01
44	98	2.4	955	8	R42248	Human p50 protein DNA	5.51e+01
45	96	2.3	422	21	W10693	Homology vector 443-8	7.31e+01

ALIGNMENTS

RESULT 1
ID W23963 standard; Protein; 593 AA.
AC W23963;
DT 17-JUN-1998 (first entry)

DE Arabidopsis thaliana NIM1 protein.
KW NIM1; noninducible immunity; disease resistance; plants;
KW SAR gene expression.
OS Arabidopsis thaliana.
PN WO9749822-A1.
PD 31-DEC-1997.
PF 10-MAR-1997; E01218.
PR 10-JAN-1997; US-035022.
PR 21-JUN-1996; US-020272.
PR 30-AUG-1996; US-024883.
PR 13-DEC-1996; US-033177.
PR 27-DEC-1996; US-773559.
PA (NOVS) NOVARTIS AG.
PI Delaney TP, Ellis DM, Friedrich LB, Johnson JE,
PI Lawton KA, Ryals JA, Weymann K;
DR WPI; 98-077185/07.
DR N-PSDB; V04632.
PT NIM1 gene which allows activation in plant of systemic acquired
PT resistance - useful to confer broad spectrum disease resistance in
PT plants, specifically crop plants, e.g. rice, wheat, barley, rye and
PT corn
PS Disclosure; Fig 15; 153pp; English.
CC The sequence is that of the NIM1 (noninducible immunity) protein. It
CC may be used to confer a broad spectrum disease resistance in plants,
CC specifically crop plants, e.g. rice, wheat, barley, rye and corn.
CC The NIM1 gene can be used to confer universal disease susceptibility
CC to plant cells, and their progeny. It can also be used in a screening
CC method for identifying compounds capable of inducing broad spectrum
CC disease resistance in plants, while the plant cells, and their
CC progeny, can be used to isolate a gene fragment which allows
CC expression of broad spectrum disease resistance in plants, or to
CC incorporate the resistant trait into plant lines through breeding.
SQ Sequence 593 AA;

Query Match 53.1%; Score 2184; DB 29; Length 593;
Best Local Similarity 51.2%; Pred. No. 3.65e-174;
Matches 307; Conservative 150; Mismatches 123; Indels 20; Gaps 14;

Db 1 mdttidgfadsyeisstsfvatdntdssivylaeeqvltpdvsalqlsnsfesvfd-s 59
|| : :|:|| :||::| :
Qy 1 MDNSRTAFSDSNDISGSSSICCIGGGMTE-FFSPE-T-SPAEITSLKRLSETLESIFDAS 57

Db 60 -pd-dfydsaklvsldg-revsfhrcvlsarssffksalaaakkekdsnntaavkleke 116
|: |:::|||||:| :|::| |||:|||||:|||:| :| |||:| :| :|:|||:
Qy 58 LPEFDYFADAKLVVSGPCKEIPVHRCILSARSPFFKN-LFCGKKEKNS---S-KVELKE 111

Db 117 iakdyevgfdsvvtvlavyssrvrpppkvsecadencchvacrpavdfmlevlylafi 176
: |: ||::|:|:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:
Qy 112 VMKEHEVSYDAVMSVLAYLYSGKVRPSPKDVCVCVDNDCSHVACRPAVALVEVLYTSFT 171

Db 177 fkipelitlyqrh1ldvvdkvviedtlvilklanicgkacmk1ldrckeiiivksnvdms 236
| |:||: :|||||:|| :| ::| :|||:|||:|||:|||:|||:|||:|||:
Qy 172 FQISELVDKFQRHLLDILDKTAADDVMMVLSVANICGKACERLLSSCIEIIVKSNVDIIT 231

Db 237 lekslpeelvkeiidrrkelglevpkvk----khvsnvhkal dsddielvkllkedhtn 292
|:|:|| :||:| | | ||||:| : | || :|||:||:||:| :||:| ||
Qy 232 LDKALPHDIVKQITDSRAELGLQGPESNGFPDKHVKRIHRALDSDDVELLQMLLREGHTT 291

Db 293 lddacalhfavaycnvktatdlkldladvnhrnprgytvlhvaamrkepqlilsllekg 352
||||| |||:|||||: |||::|| | |||:||:|||:||||| :|||:|||
Qy 292 LDDAYALHYAVAYCDAKTTAELLDLALADINHQNSRGYTVLHVAAMRKEPKIVVSLLTKG 351

Db 353 asaseatlegrtalmiakqatmavecn ni peqckhs lkgrlcveileqedkreqiprdvp 412
| :|: | :|| | |||: | | :|:||: | : :| |||:|||: :||: : :
Qy 352 ARPSDLTSDGRKALQIAKRLTRLVDFSKSPEEGKSASNDRLCIEILEQAERRDPLLGEAS 411

Db 413 psfavaadelkmtlldlenrvlaqrlfp teaqaaameiaemkgtcefivtslepdrltgt 472
|:|:|:|:|:| | |||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:
Qy 412 VSLAMAGDDL RMKLLYLENRVGLAKLLFPM EAKVAMDIAQVDGTSEFPLASIG-KKMANA 470

Db 473 krtspgvkiapfrileehqsrlkalsktvelgkrffprcsavldqimncedltqlac ed 532
||: :| |||:| | |||:|||:|||:|||:|||:|||:|||:|||:|||:|||:
Qy 471 QRTTVDLNEAPFKIKEEHLNRLRALSRTVELGKRFFPRCSEVLNKIMDADDLSEIAYMGN 530

Db 533 dtaekrlqkkqrymeiqetlkkaf sednle gnss ltdst ss ts kstggkrs nrk lsh rr 592
|||||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:
Qy 531 DTAEERQLKKQRYMELQEILTKAFTEDKEEYDKTN-NI SSSCSSTS KGVDKPN-KLPFRK 588